Advances in Digestive Medicine (2015) 2, 41—43

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.aidm-online.com

EDITORIAL

Analyzing Taiwan's National Health Insurance Research Database to explicate the allocation of health-care resources

More than 99% of the residents in Taiwan have been covered by National Health Insurance (NHI) since its implementation in 1995, thanks to the single-payer, universal, and compulsory policies [1]. To construct and maintain a publicly available database for the purpose of academic research, the government entrusts the Taiwan National Health Research Institute (NHRI) to continuously and systemically collect relevant registration and claim data generated in the insurance system. Therefore, the National Health Insurance Research Database (NHIRD) contains comprehensive computerized NHI records of literally the entire population of Taiwan for the past 2 decades. Under the "least privilege" principle, qualified researchers are entitled to access the database after their projects are reviewed and approved by the NHRI [2]. As a reliable data source, the NHIRD has facilitated academic research in various scientific disciplines, and the rapidly increasing number of research articles is proof of this [3].

In addition to imposing a low barrier for application, the NHIRD has several advantages that continue to attract investigators. The almost 100% coverage is an obvious strength that enables researchers to test their hypotheses on the scale of a nationwide population. This is a unique feature that distinguishes the NHIRD from many other electronic health-care databases, which either register a specific population (e.g., the Veterans Administration or Medicare databases from the United States) or comprise selected samples (e.g., Clinical Practice Research Datalink from the United Kingdom) [4—6]. Because the NHRI continuously collects data, and annually updates the database, investigators are able to longitudinally track a beneficiary from as early as 1995 to as recent as the preceding year of study approval. These advantages greatly enhance statistical power when studying rare diseases or infrequent outcomes [7]. Moreover, the NHIRD contains information about every billing order, including details on service type, claiming doctor (anonymously coded), product brand, prescription date, and the cost. Researchers can figure out where, when, how much, how frequently, and for how long a patient received a specified intervention [8]. Even different brands of a generic drug are identifiable [9]. Another easily overlooked strength of the NHIRD is the unbiased process of data collection. This merit minimizes the concern of observer-expectancy effect, given that the information available is not collected for a specific study. For investigators' convenience, the NHRI issues the Longitudinal Health Insurance Database (LHID) as a representative NHIRD subset. The LHID is composed of details on 1 million individuals randomly sampled from all beneficiaries. Its representativeness is supported by similar distributions in age, sex, and health-care cost from the original NHIRD.

However, the NHIRD has crucial limitations that investigators have to consider when designing their study [10]. First, it contains little information on data that are not necessary for reimbursement claims. Lacking examination results should be recognized as a major weak point. Similarly, only limited information is available on socioeconomic status, lifestyle, health behavior, or unhealthy habits. Accordingly, researchers should carefully define the diseases and outcomes, bearing in mind that the diagnostic code in the NHIRD is principally meant for billing purposes and may not be accurate enough by itself. There are several ways to overcome this limitation. If applicable, the Registry for Catastrophic Illness Patient Database can be searched to secure a reliable definition [11]. Entry into this registry, which waives copayment of the enrollee, is permitted only after explicit criteria are fulfilled. The primary diagnoses for hospitalization, particularly those for a severe illness such as ischemic stroke or acute myocardial infarction [9,12,13], are also convincing in general. With regard to outpatient condition or underlying comorbidity, additional prerequisites other than the diagnostic code can enhance the reliability. For instance, pharmacotherapy that defines the diseases is useful [14]. Another major

http://dx.doi.org/10.1016/j.aidm.2015.04.001

limitation of the NHIRD is the strict regulations that prohibit investigators from linking the de-identified database with external information, for the sake of privacy protection. Linkage with other public registry databases is possible but cannot be performed outside the designated offices which are tightly controlled by the government. Furthermore, the basic unit that forms the NHIRD is details of an individual person but without information that can pinpoint kinship or blood relationship. This is undoubtedly an impassable obstacle when exploring inheritable traits or household situations. Finally, the NHIRD includes no data on the self-pay or trial settings. This limitation, fortunately, is not a critical issue under most circumstances, because health services encompassed in the NHI range extensively from preventive measures to hospice care.

In this issue of the *Journal*, Yeh and colleagues [15] added another example to the growing bulk of literature brought about by analyzing the NHIRD. They utilized data from LHID (2005), which contains 1 million random samples, to study the expenditure associated with liver cancer. A total of 2335 patients with a specific diagnostic code were identified and their reimbursement claims in 2009 were analyzed. Perhaps to no one's surprise, liver cancer patients utilized more NHI resources compared with those without the diagnosis, at the ratio of 666.8% for each patient for the year. The authors also categorized utilization according to the type of service. Intriguingly, sex, age, income status, residential region, and insured classification seemed to influence how much a patient was reimbursed. These investigators are to be commended for reminding us how liver cancer, most of which is identified as hepatocellular carcinoma in histopathologic analysis, has been financially devastating our country. As a leading cause of cancer mortality for decades, liver cancer is expensive not only because of frequent hospital visits, but also because of the costly services required for diagnosis and treatment. Given that the incidence of liver cancer remains high and new therapies keeps hitting the market (e.g., sorafenib was not covered in Taiwan until 2012), this malignant disease most likely will continue to take a big bite from the global budget of NHI.

Yeh et al.'s [15] study, nonetheless, had a few caveats that warrant attention. The inability to estimate expenditure outside the insurance system is a major limitation. Conceivably, patients with malignant disease are prone to seek alternative or folk medicine. In addition, indirect financial loss was not studied. Another NHIRD-inherent limitation is the difficulty to ascertain socioeconomic information. Although their study resorted to insured classification and premium, the accuracy of these parameters to stand for income status was not grounded. Moreover, their investigation was mainly descriptive rather than analytical. The apparent association of medical expenditure with age, sex, residential region, and insured classification was regrettably not further scrutinized. Similarly, the influence of comorbidity was not addressed. Finally, this study would have achieved a greater impact if the authors had considered patient outcomes. Undoubtedly, liver cancer treatment is costly, but what would be more informative is to elucidate why it is so expensive and how the costs are associated with clinical outcomes and care quality.

The NHIRD is going to deliver more and more research in the foreseeable future. With such an abundance of claim data recorded in detail, it provides an appealing venue for studying health economics, care quality, and cost effectiveness. We now know liver cancer is disproportionately consuming national health-care resources, and hopefully we will soon learn more details explicating the expense allocated to manage this both deadly and costly disease.

Conflicts of interest

The author declares no conflicts of interest.

References

[1] Cheng TM. Taiwan's new National Health Insurance program: genesis and experience so far. Health Aff (Millwood) 2003;22:61—76.
[2] Offical website of National Health Insurance Research Database: Available at: http://nhird.nhri.org.tw/en/index.htm. [last accessed: February 2015].
[3] Chen YC, Wu JC, Chen TJ, Wetter T. Reduced access to database. A publicly available database accelerates academic production. BMJ 2011;342:d637.
[4] Chen YC, Wu JC, Haschler I, Majeed A, Chen TJ, Wetter T. Academic impact of a public electronic health database: bibliometric analysis of studies using the general practice research database. PloS One 2011;6:e21404.
[5] Hynes DM, Cowper D, Kerr M, Kubal J, Murphy PA. Database and informatics support for QUERI: current systems and future needs. Quality Enhancement Research Initiative. Med Care 2000;38:I114—28.
[6] Warren JL, Klabunde CN, Schrag D, Bach PB, Riley GF. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Med Care 2002;40:IV-3—18.
[7] Hsu YC, Ho HJ, Huang YT, Wang HH, Wu MS, Lin JT, et al. Association between antiviral treatment and extrahepatic outcomes in patients with hepatitis C virus infection. Gut 2015;64:495—503.
[8] Hsu YC, Ho HJ, Wu MS, Lin JT, Wu CY. Postoperative peg-interferon plus ribavirin is associated with reduced recurrence of hepatitis C virus-related hepatocellular carcinoma. Hepatology 2013;58:150—7.
[9] Wu CY, Chan FK, Wu MS, Kuo KN, Wang CB, Tsao CR, et al. Histamine2-receptor antagonists are an alternative to proton pump inhibitor in patients receiving clopidogrel. Gastroenterology 2010;139:1165—71.
[10] Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323—37.
[11] Hsu YC, Lin JT, Chen TT, Wu MS, Wu CY. Long-term risk of recurrent peptic ulcer bleeding in patients with liver cirrhosis: a 10-year nationwide cohort study. Hepatology 2012;56:698—705.
[12] Cheng CL, Lee CH, Chen PS, Li YH, Lin SJ, Yang YH. Validation of acute myocardial infarction cases in the National Health Insurance Research Database in Taiwan. J Epidemiol 2014;24:500—7.
[13] Cheng CL, Kao YH, Lin SJ, Lee CH, Lai ML. Validation of the National Health Insurance Research Database with ischemic stroke cases in Taiwan. Pharmacoepidemiol Drug Saf 2011;20:236—42.
[14] Hsu YC, Lin JT, Ho HJ, Kao YH, Huang YT, Hsiao NW, et al. Antiviral treatment for hepatitis C virus infection is associated

with improved renal and cardiovascular outcomes in diabetic patients. Hepatology 2014;59:1293—302.
[15] Yeh C, Lin Y-H, Yu C-P, Hsu C-T. Medical utilization by liver cancer patients under the National Health Insurance program in Taiwan: a population-based cross-sectional study. Adv Dig Med 2015;2:46—53.

Yao-Chun Hsu*
School of Medicine for International Students, I-Shou University, Kaohsiung, Taiwan

Division of Gastroenterology and Hepatology, Department of Internal Medicine, E-Da Hospital/I-Shou University, Kaohsiung, Taiwan

*Center for Database Research and Department of Internal Medicine, E-Da Hospital/I-Shou University, Number 1, E-Da Road, Kaohsiung 824, Taiwan.
E-mail address: holdenhsu@gmail.com